ORIGINAL ARTICLE

Controlling the Flow of Distracting Information in Working Memory

Nicole Hakim [1,2], Tobias Feldmann-Wüstefeld[3], Edward Awh[1,2,4] and Edward K. Vogel[1,2,4]

[1]Department of Psychology, University of Chicago, Chicago, IL 60637, USA, [2]Institute for Mind and Biology, University of Chicago, Chicago, IL 60637, USA, [3]Department of Psychology, University of Southampton, Southampton SO17 1PS, UK and [4]Grossman Institute for Neuroscience, Quantitative Biology, and Human Behavior, University of Chicago, Chicago, IL 60637, USA

Address correspondence to Nicole Hakim, 940 E 57th Street, Chicago, IL 60637, USA. Email: nhakim@uchicago.edu.

Abstract

Visual working memory (WM) must maintain relevant information, despite the constant influx of both relevant and irrelevant information. Attentional control mechanisms help determine which of this new information gets access to our capacity-limited WM system. Previous work has treated attentional control as a monolithic process—either distractors capture attention or they are suppressed. Here, we provide evidence that attentional capture may instead be broken down into at least two distinct subcomponent processes: (1) Spatial capture, which refers to when spatial attention shifts towards the location of irrelevant stimuli and (2) item-based capture, which refers to when item-based WM representations of irrelevant stimuli are formed. To dissociate these two subcomponent processes of attentional capture, we utilized a series of electroencephalography components that track WM maintenance (contralateral delay activity), suppression (distractor positivity), item individuation (N2pc), and spatial attention (lateralized alpha power). We show that new, relevant information (i.e., a task-relevant distractor) triggers both spatial and item-based capture. Irrelevant distractors, however, only trigger spatial capture from which ongoing WM representations can recover more easily. This fractionation of attentional capture into distinct subcomponent processes provides a refined framework for understanding how distracting stimuli affect attention and WM.

Key words: distraction, interruption, suppression, working memory gating

Introduction

A critical feature of working memory (WM) is to protect internal representations from external interference, For example, when driving, the WM representation of our route must be maintained despite irrelevant external interference, such as a flashing colorful billboard. Nevertheless, external information is sometimes relevant, and WM must integrate this new information with our ongoing WM representations. For example, a flashing sign warning us of a car accident ahead may capture our attention— but to our advantage. This new information allows us to update our WM representation of our route in order to avoid traffic caused by the car accident. Attentional control mechanisms help determine which information gets access to our capacity-limited WM system.

Models of attentional control suggest that attentional selection is based on a competitive process (Duncan and Humphreys 1989; Bundesen 1990) in which both goal-driven and stimulus-driven factors determine which information is selected from a given display (Wolfe and Horowitz 2004; Posner et al. 2004; Jonides and Irwin 1981; Posner and Petersen 1990; Folk et al. 1992; Desimone and Duncan 1995; Egeth and Yantis 1997; Itti and Koch 2000; Kastner and Ungerleider 2001; Corbetta and Shulman 2002; Awh et al. 2012). However, can goal-driven attentional selection override stimulus-driven capture? Some previous research has

suggested that salient irrelevant distractors capture attention before it can be suppressed (Hickey et al. 2009; Sawaki et al. 2012; Feldmann-Wüstefeld and Schubö 2013; Liesefeld et al. 2017). Based on this, they argue that attentional capture is obligatory because salient information needs to be processed before being discarded. Other research suggests that distractor suppression can prevent attentional capture (Gaspar and McDonald 2014; Gaspelin et al. 2015, 2017; Gaspelin and Luck 2018; Feldmann-Wüstefeld et al. 2020). This work has found that, for example, participants make fewer erroneous eye movements towards (Gaspelin et al. 2017) and are less likely to report the identity of (Gaspelin et al. 2015) a successfully suppressed salient distractor than a less salient distractor that was not suppressed. Although these appear to be conflicting perspectives, we suggest here that they might be reconciled by distinguishing between two distinct forms of attentional capture.

Attentional capture and suppression are often treated as a monolithic process: The onset of a stimulus results in either capture or suppression of the novel information. However, recent work suggests that attention may include two distinct subcomponent processes: attention to regions in space and representations of objects that occupy the attended regions (Prinzmetal et al. 2009; Zivony and Lamy 2018; Hakim et al. 2019a; Maxwell et al. 2020). In line with these studies, we propose that involuntary attentional capture may also be broken down into at least two distinct subcomponent processes: (1) Spatial capture, which refers to when spatial attention shifts towards the location of irrelevant stimuli and (2) item-based capture, which refers to when item-based representations of distracting stimuli are formed in WM. In the current work, we obtained electroencephalography (EEG) evidence to directly test whether spatial capture is distinct from item-based capture.

We separately measured spatial and item-based capture using EEG markers of spatial and item-based attentional capture. We used lateralized alpha power (8–12 Hz) to track spatial capture (Foster et al. 2016; Hakim et al. 2019a). This oscillatory signal has been shown to track attended hemifield (Hakim et al. 2019a) and has been shown to contain precise spatial information about attended stimuli (Foster et al. 2016, 2017) (We would like to note that there is a strong consensus that the locus of attention is reflected in the topography of alpha power (Foster et al. 2016, 2017; Hakim et al. 2019). However, further research is needed to understand the complex interrelationship between alpha power and object-related representations (see alpha power results from Hakim et al. 2019a)). To track item-based capture, we used the contralateral delay activity (CDA) and the distractor positivity (P_D). The CDA tracks the number of items maintained in WM (Luria et al. 2016; Balaban and Luria 2017), whereas the P_D tracks the suppression of irrelevant information (Hickey et al. 2009; Feldmann-Wüstefeld and Schubö 2013; Burra and Kerzel 2014), while also being sensitive to the number of irrelevant items that are presented (Feldmann-Wüstefeld and Vogel 2018).

We used these EEG signals to assess how salient items with a sudden onset (distractors) are processed when subjects are maintaining relevant information in WM. In addition, we manipulated the task relevance of the new, distracting information to compare how each type of attentional capture is influenced by goal-driven selection. Finally, although past work has typically focused on competition between simultaneously presented targets and distractors, here we focused on how task-relevant and task-irrelevant distractors (We would like to note that previous

work has used the term "distractor" to refer to task-irrelevant information and "interrupter" to refer to task-relevant information (Trafton et al. 2003; Clapp and Gazzaley 2012). Here, we chose to use the terms "task-relevant distractors" and "task-irrelevant distractors" in order to emphasize our main experimental manipulation of relevance. This terminology also aligns with the existing attention capture literature (Spinks et al. 2004; Olivers et al. 2006; Hollingworth and Beck 2016).) influence the maintenance of items that have already been stably encoded into WM. This provided the opportunity to obtain clear evidence regarding the degree to which distractors elicited spatial and item-based attentional capture, and the distinct impact of task relevance on each form of capture.

Prominent models of attentional control assert that visually selected stimuli should automatically gain access to WM at least for a short period of time (Bundesen et al. 2005). Moreover, previous research has shown that a sudden onset of salient but irrelevant information captures attention (Yantis and Jonides 1984; Franconeri and Simons 2003; Theeuwes 2010; Feldmann-Wüstefeld et al. 2015) and that this negatively impacts ongoing WM representations (Bisley and Goldberg 2010; van Moorselaar et al. 2017; Hakim et al. 2019b). Does this distraction of WM maintenance reflect an obligatory encoding of this new information into WM? To anticipate the results, we observed clearly distinct effects of task relevance on spatial and item-based attentional capture. Continuous tracking of alpha laterality showed that spatial attention was captured by distracting stimuli, regardless of whether they were relevant or not. In sharp contrast, item-based attentional capture was completely determined by task relevance. Task-relevant distractors were encoded into WM, as shown by N2-posterior-contralateral (N2pc) and CDA signals that tracked item individuation and WM maintenance, respectively. By contrast, a P_D was observed contralateral to the task-irrelevant distractors and no CDA was observed. Thus, task-irrelevant distractors were not actively encoded into WM, even though they clearly captured spatial attention. Thus, our findings offer a potential reconciliation of prior conflicting findings by showing that the encoding of distracting information into WM can be suppressed even when there is clear evidence that spatial attention has been captured.

Materials and Methods

Experiment 1

In Experiment 1, we sought to determine how task-relevant versus task-irrelevant distractors are processed. To this end, we presented memory array items along the midline and presented distractors laterally. This allowed us to isolate the neural representations of the distractors themselves. With this design, any lateralized signal, such as CDA or lateralized alpha power, should reflect the processing of the distractors and not the memory array.

Previous research has shown that active representations may be required for the identification of relevant stimulus features (Mazza et al. 2007; Mcdonald et al. 2013). Therefore, we predicted that when participants had to discriminate the task-relevant distractors (discriminate condition), they would be more likely to encode them into WM than when task-irrelevant distractors were presented (ignore condition). Accordingly, there should be a CDA following task-relevant, but not following task-irrelevant distractors. Conversely, when participants could ignore the

task-irrelevant distractors, we predicted that they would actively suppress them, as their features do not need to be identified. The P_D has been shown to track suppression of irrelevant information (Sawaki and Luck 2012; Feldmann-Wüstefeld and Schubö 2013; Burra and Kerzel 2014). Therefore, we should expect to find a robust P_D when participants ignore the task-irrelevant distractors, but not when they discriminate the task-relevant distractors.

Participants

A total of 30 novel volunteers, naïve to the objective of the experiment, participated for payment ($15 USD per hour). Data from one participant was excluded from the analysis because of technical issue with the behavioral data file. Data from nine participants were excluded from the analysis because of too many artifacts that resulted in fewer than 150 trials in any condition. The remaining 20 participants (six male) were between the ages of 21 and 31 [M = 23.5, standard deviation (SD) = 3.3]. Participants in all experiments reported normal or corrected-to-normal visual acuity as well as normal color vision. All experiments were conducted with the written understanding and consent of each participant. The University of Chicago Institutional Review Board approved experimental procedures.

Stimuli

All stimuli were presented on a gray background ($\sim$33.3 cd/m^2). Cue displays showed a central fixation dot (0.2° × 0.2°). Memory displays showed four colored squares (1.1° × 1.1°, mean luminance 43.1 cd/m^2) along the midline with a randomly jittered horizontal offset of maximally 0.55° (half of an object). Colors for the squares were selected randomly from a set of 11 possible colors (red = 255, 0, 0; green = 0, 255, 0; blue = 0, 0, 255; yellow = 255, 255, 0; magenta = 255, 0, 255; cyan = 0, 255, 255; purple = 102, 0, 102; brown = 102, 51, 0; orange = 255, 128, 0; white = 255, 255, 255; black = 0, 0, 0). No color was repeated. On 50% of trials, the retention interval display remained blank with a central fixation dot (0.2° × 0.2°). However, on the other 50% of trials, during the delay the distracting stimuli appeared laterally. On one side of the screen, four colored circles (25% of all trials) or four squares (25% of all trials) appeared during the delay. Items from these colors were chosen from the say 11 possible colors, but were never the same as the memory array items on a given trial. In addition, these items had the same area as the items from the memory array. On the other side of the screen, four gray diamonds appeared [red, green, blue (RGB) = 80, 80, 80] at the same time as the colored circles/squares. These gray diamonds were the same area as the colored circles/squares. The gray diamonds were also luminance matched (i.e., iso-luminant) to the average of the colored circles/squares. We presented these gray diamonds so as to match the bottom-up visual stimulation on both sides of the screen. The hemifield in which the diamonds and the hemifield in which the colored circles/squares were presented were randomly selected in each trial. All stimuli had the same area. Probe displays showed one colored square along the midline in the same location as one of the memory array items, randomly picked, in the original array. In 50% of the trials, the color of the square in the attended hemifield was identical (no change trial) to the memory display. In the remaining 50% of trials, it was one of the colors not used in the memory or distractor display (change trials).

Apparatus

Participants were seated with a chin rest in a comfortable chair in a dimly lit, electrically shielded and sound attenuated chamber. Participants responded with button presses on a standard keyboard that was placed in front of them. Stimuli were presented on an LCD computer screen (BenQ XL2430T; 120 Hz refresh rate; 61 cm screen size in diameter; 1920 × 1080 pixels) placed at 74 cm distance from participants. An IBM-compatible computer (Dell Optiplex 9020) controlled stimulus presentation and response collection.

Procedure

Each trial began a memory display consisting of four colored squares along the midline appeared for 150 ms. Participants were instructed to memorize as many colored squares in the memory display as possible. Participants had to remember the items over a blank retention interval that contained a central fixation dot. The retention interval lasted 2000 ms, regardless of whether distractors appeared. In 50% of the trials, a distractor display appeared 500 ms after memory display offset for 150 ms. The distracting display consisted of four circles (50% of distractor trials) or four squares (50% of distractor trials) that appeared laterally. This display was visually balanced by four iso-luminant gray diamonds that appeared in the opposite hemifield. In half of the trials, participants were instructed to ignore the distractor displays (Ignore block). In the other half of the trials, participants were instructed to discriminate the shape of the distractor items (Discriminate block). After the retention interval, a probe display appeared until response. In both Ignore and Discriminate blocks, participants had to indicate whether the object at the probed location of the attended hemifield changed color ("?/" key) or did not change color ("z" key). In the Discriminate block, participants additionally performed a go-no-go task. Before responding to the probe, they had to press "space" to indicate that the distractor objects were circles. If the distractor objects were squares, they did not have to press any key. After participants responded, the next trial started after a blank intertrial interval of 750 ms. Participants completed a total of 1600 trials (20 blocks of 80 trials), that is, 800 trials with distractors and 800 trials without distractors. The first half of the experiment was always the Ignore blocks (task-irrelevant distractors), and the second half of the experiment was always the Discriminate blocks (task-relevant distractors). Information about average performance and a minimum break of 30 s was provided after each block. See Figure 1 for a visual depiction of the task.

We presented the distractors in locations that did not overlap with the locations of the memory items to avoid visual masking. Importantly, the relative position of distractors and targets matter in this kind of change detection tasks. When distractors are presented laterally with targets on the vertical midline, the neural signature of sustained distractor suppression can be isolated (CDAp). Conversely, when distractors are presented on the vertical midline and targets are presented laterally, the neural signature of target processing can be isolated (Feldmann-Wüstefeld and Vogel 2018). Accordingly, in Experiment 1, we were interested in the neural representations of the distractors, so we placed the distractors laterally. Thus, lateralized signals, such as CDA and lateralized alpha power, could be used to assess item-based and spatial capture elicited by the lateralized distractors.
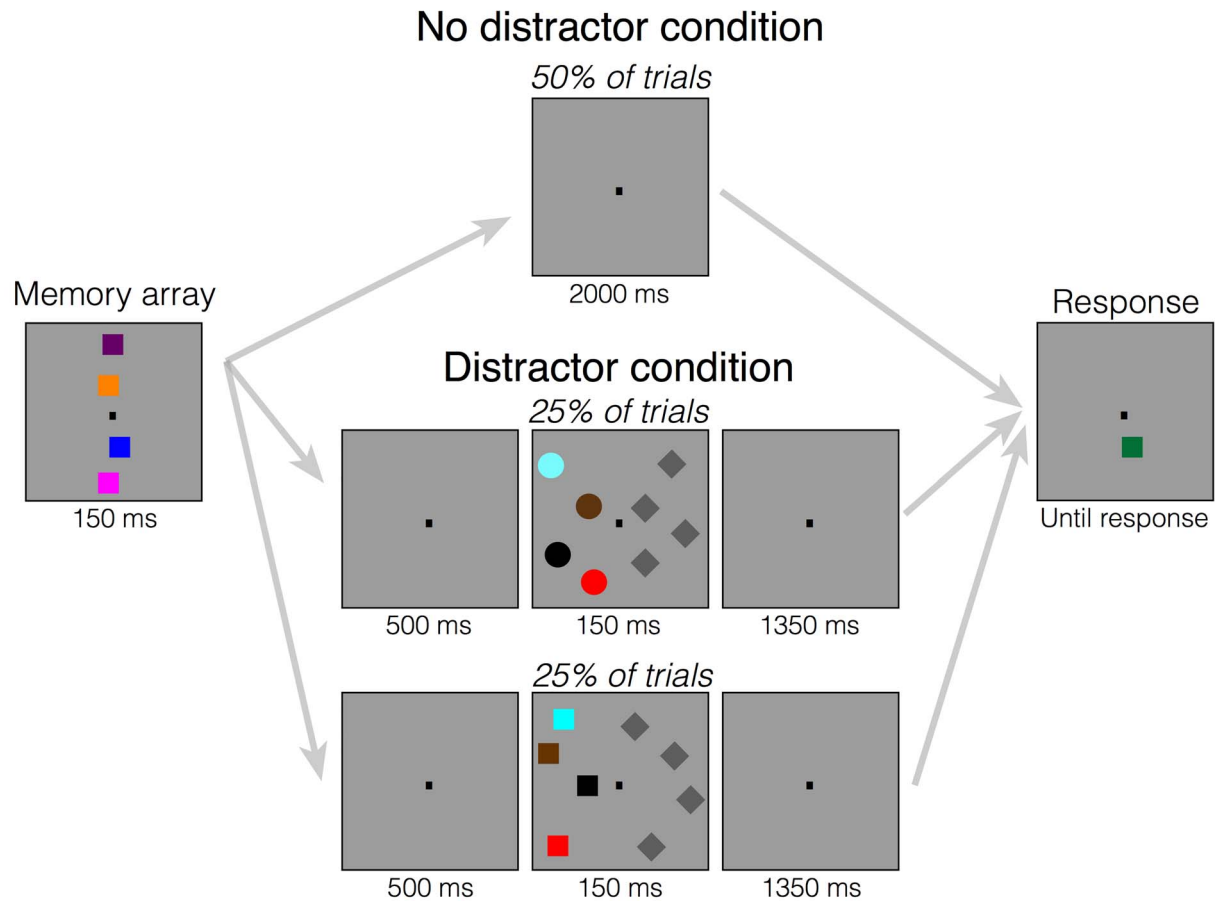
Figure 1. Task design for Experiment 1. At the start of each trial, the memory array appeared, which consisted of four colored squares along the midline. Participants were told to remember the colors and locations of these squares over the brief delay. Following the memory array, the screen went blank. Then, either the screen remained blank the entire delay (no distractor condition) or the screen remained blank for 500 ms (distractor condition) followed by a series of distractors presented laterally. This distractor array consisted of either four colored squares or circles on one side of the screen and iso-luminant gray diamonds on the other side of the screen. When participants were in the "Ignore" condition, they were told to always ignore these task-irrelevant distractor objects. When they were in the "Discriminate" condition, they were told to determine the shape of the colored stimuli (squares vs. circles) in order to report whether the stimuli were circles. They were told to withhold their response until the response screen appeared. Following the distractor array, the screen then went blank for the rest of the delay. On the final screen, one square on the midline reappeared and could either be the same color as the original square or it could be a different color. In both conditions, participants had to report whether the square on the cued side of the screen changed colors. In the "Discriminate" condition, participants additionally had to report whether the distractor were circles, if there were distractor on that trial.

Artifact rejection

We recorded EEG activity from 30 active Ag/AgCl electrodes (Brain Products actiCHamp) mounted in an elastic cap positioned according to the International 10–20 system (Fp1, Fp2, F7, F8, F3, F4, Fz, FC5, FC6, FC1, FC2, C3, C4, Cz, CP5, CP6, CP1, CP2, P7, P8, P3, P4, Pz, PO7, PO8, PO3, PO4, O1, O2, Oz). FPz served as the ground electrode and all electrodes were referenced online to TP10 and re-referenced offline to the average of all electrodes. Incoming data were filtered (low cutoff = 0.01 Hz, high cutoff = 80 Hz, slope from low- to high-cutoff = 12 dB/octave) and recorded with a 500 Hz sampling rate. Impedances were kept below 10 kΩ. To identify trials that were contaminated with eye movements and blinks, we used electrooculogram (EOG) activity and eye tracking. We collected EOG data with five passive Ag/AgCl electrodes (two vertical EOG electrodes placed above and below the right eye, two horizontal EOG electrodes placed ~1 cm from the outer canthi, and one

ground electrode placed on the left cheek). We collected eye-tracking data using a desk-mounted EyeLink 1000 Plus eye-tracking camera (SR Research Ltd) sampling at 1000 Hz. Usable eye-tracking data were acquired for 20 out of 22 participants in Experiment 1 and 29 out of 30 participants in Experiment 2.

EEG was segmented offline with 2000 ms segments time-locked to memory display onset, including a 200 ms prestimulus baseline period. Eye movements, blinks, blocking, drift, and muscle artifacts were first detected by applying automatic detection criteria to each segment. After automatic detection, trials were manually inspected to confirm that detection thresholds were working as expected. Incorrect trials and any trial contaminated with artifacts were excluded from analysis. For example, a trial including a task-relevant distraction was only considered as "correct" and included in further analyses if both the response to the primary and the interrupting task were correct. The removal of all trials with artifact(s) allowed us to ensure that

there was not any missing data within any of the included trial epochs.

For the participants used in analyses, we rejected on average 12.0% (SD = 12.7%) of trials in Experiment 1 and 27.0% (SD = 6.3%) of trials in Experiment 2 per person. Collapsed across conditions, this resulted in an average of 1408 remaining trials per person in Experiment 1 and 1168 remaining trials in Experiment 2. In order to achieve an acceptable signal-to-noise ratio, participants were excluded if fewer than 150 correct trials were available in any of the conditions. Collapsed across conditions, participants in Experiment 1 were correct on 79.9% of all trials, and in Experiment 2, they were correct on 78.7% of trials. Broken down by condition, participants in Experiment 1 were correct on 82.4% of trials without distractors and 77.3% of distractor-present trials. In Experiment 2, they were correct on 83.2% of trials without distractors and 74.3% of distractor-present trials.

Eye Movements

We used a sliding window step function to check for eye movements in the horizontal electrooculogram (HEOG) and the eye-tracking gaze coordinates. For HEOG rejection, we used a split-half sliding window approach. We slid a 100 ms time window in steps of 10 ms from the beginning to the end of the trial. If the change in voltage from the first half to the second half of the window >20 μV, it was marked as an eye movement and rejected. For eye-tracking rejection, we applied a sliding window analysis to the x-gaze coordinates and y-gaze coordinates (window size = 100 ms, step size = 10 ms, threshold = 0.5° of visual angle).

Blinks

We used a sliding window step function to check for blinks in the vertical electrooculogram (VEOG) (window size = 80 ms, step size = 10 ms, threshold = 30 μV). We checked the eye-tracking data for trial segments with missing data points (no position data is recorded when the eye is closed).

Drift, Muscle Artifacts, and Blocking

We checked for drift (e.g., skin potentials) by comparing the absolute change in voltage from the first quarter of the trial to the last quarter of the trial. If the change in voltage exceeded 100 μV, the trial was rejected for drift. In addition to slow drift, we checked for sudden step-like changes in voltage with a sliding window (window size = 100 ms, step size = 10 ms, threshold = 100 μV). We excluded trials for muscle artifacts if any electrode had peak-to-peak amplitude >200 μV within a 15 ms time window. We excluded trials for blocking if any electrode had at least 30 time points in any given 200 ms time window that were within 1 V of each other.

Behavioral Data Analysis

We separately analyzed performance for four separate conditions: trials without distractors in the ignore block, trials with distractors in the ignore block (task-irrelevant distractors), trials without distractors in the discriminate block, and trials with distractors (task-relevant distractors) in the discriminate block. Performance was converted to a capacity score, K, calculated as $N \times (H\text{-FA})$, where N is the set size, H is the hit rate, and FA is the false alarm rate (Cowan 2001). To compare performance between conditions, we used a two-way analysis of variance (ANOVA) with the within-subjects factors Distraction (relevant vs. irrelevant) and Relevance (ignore vs. discriminate).

All analyses were done with circle and square distractors collapsed, since the circles and squares were equiprobable (circles: 50% of distractor trials, squares: 50% of distractor trials). We additionally ran two-tailed follow-up t-tests when it was justified.

Lateralized event-related potential (ERP) Analyses

Segmented EEG data was baselined from 200 to 0 ms before the onset of the memory displays. Artifact-free EEG segments were averaged separately for the two conditions when distractors appeared (irrelevant vs. relevant). Data was not analyzed for trials without distractors because "laterality" was undefined in this condition. The difference between contralateral and ipsilateral activity for the electrode pair PO7/PO8 was calculated (i.e., the CDA), resulting in two average waveforms for each participant (one per analyzed condition). The average CDA amplitude was calculated for three time windows: before distractor onset (400–650 ms), and two windows following distractor offset (850–950 ms and 1050–1300 ms). Previous research has shown that CDA amplitude should stabilize ~400 ms after memory array onset. To measure CDA after it is stabilized, we chose a time window starting 400 ms after onset of the memory array. We wanted to measure an analogous time window following the onset of the distractors (i.e., 400 ms after distractor onset), which is why we chose the time window 1050–1300 ms. For all of these time windows, we then compared the CDA across conditions with a paired samples t-test. To measure the robustness of the CDA for each condition (reliable difference between contra- and ipsilateral activity), we also ran t-tests (against zero) for each time window and condition. These t-tests are two-tailed, unless otherwise stated. We corrected for multiple comparisons using a Bonferroni correction. We applied this correction to the two postdistractor time windows. Thus, P values <0.025 from these two postdistractor time windows are considered significant. The predistractor time window (400–650 ms) was not included in our correction for multiple comparison because it is logically impossible for there to be differences across conditions in this time window because participants were not able to predict when they would be distracted on any given trial. We analyzed this time window so as to obtain a measure of noise. We would also like to note that a Bonferroni correction in this case is extremely conservative, as we had strong a priori expectations about when to expect an effect in the postdistraction I (850–950 ms) time window (Feldmann-Wüstefeld and Vogel 2018). We also chose the postdistraction II (1050–1300 ms) time window before analyzing the data, though this analysis window was more exploratory.

On trials with distractors, we additionally analyzed the P_D and the N2pc. To calculate these signals, we used a data-driven approach from previous research to specify the specific time windows of interest (Feldmann-Wüstefeld and Vogel 2018). To calculate the lateralized waveform (contra- minus ipsilateral to colored distractors) for electrodes PO7/PO8, across participants and conditions, we determined the peak of the P_D and N2pc as the most positive or negative peak, respectively, 200 to 350 ms after distractor onset across both conditions. The average amplitude from 20 ms before to 20 ms after that peak was used for statistical analyses on the P_D and the average amplitude from 50 ms before to 50 ms after that peak was used for statistical analyses on the N2pc. As an exploratory analysis, we also plotted ERP topographies for both experiments (see Supplemental materials Figures 1 and 3).

Lateralized Alpha Power Analysis

For the alpha power analyses, we did not baseline the segments. The raw EEG signal was band-pass filtered in the alpha band (8–12 Hz) using a two-way least-squares finite-impulse-response filter ("eegfilt.m" from EEGLAB Toolbox). Instantaneous power was then extracted by applying a Hilbert transform ("hilbert.m") to the filtered data. The resulting data were averaged separately for the two conditions when distractors appeared (relevant vs. irrelevant distractors) and each laterality (contra- vs. ipsilateral to cued hemifield) for the electrode pair PO7/PO8. Average alpha power was calculated for two of the same time windows as the CDA analysis: before distractor onset (400–650 ms) and postdistractor offset I (850–950 ms). Previous research has not investigated lateralized alpha power while participants maintained WM representations that were presented centrally. Therefore, even though we chose our time windows before analyzing our results, in this experiment, we did not have strong a priori predictions about the timing of alpha power lateralization following the onset of lateralized distractors. In addition, even when relevant memoranda are presented laterally, alpha power takes up to 1000 ms after memory array onset to become fully lateralized (Hakim et al. 2019a, 2019b). Therefore, the third time window that we analyzed extended to included up to 1000 ms after distractor onset (1050–1650 ms). We then compared alpha power lateralization for each time window with a paired samples t-test. To measure the robustness of alpha power lateralization for each condition (reliable difference between contra- and ipsilateral activity), we also ran t-tests (against zero) for each time window and condition. These t-tests are two-tailed, unless otherwise stated. Once again, we corrected the two postdistractor time windows for multiple comparisons using a Bonferroni correction for two comparisons (significance threshold: $P < 0.025$). As an exploratory analysis, we also plotted the topography of alpha power across the entire scalp (see Supplemental Materials Figures 2 and 4).
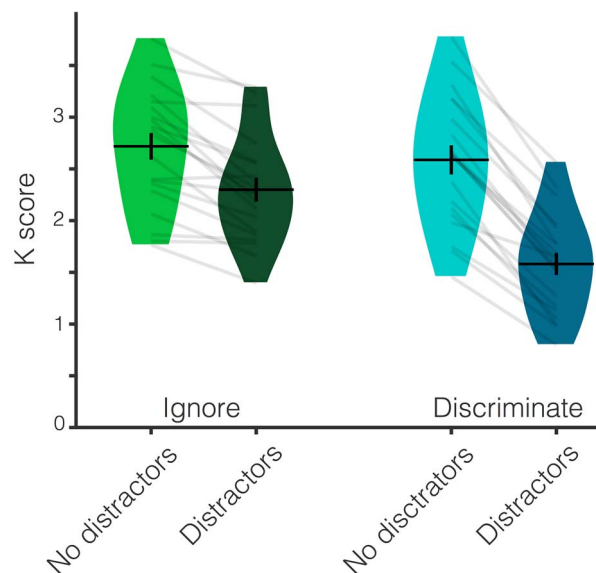
Results

Behavior

Behavioral performance was significantly above chance in all conditions (all one-sample t-tests, $P \leq 0.001$). Participants remembered fewer items when they were distracted (M = 1.95, SD = 0.61) than when they were not distracted (M = 2.66, SD = 0.60). This difference was larger when the distractors had to be discriminated (ΔM = 1.001, SD = 0.297) than when they could be ignored (ΔM = 0.419, SD = 0.293). This was evident from the significant interaction of Distraction and Relevance ($F_{(1,19)} = 55.725$, $P < 0.001$, $\eta_p^2 = 0.746$) and from the significant follow-up paired samples t-test ($t(19) = -7.465$, $P < 0.001$). This t-test compared the difference between trials with and without distractors in the Ignore and the Discriminate conditions. The main effects of Distraction and Relevance were also significant (both $P < 0.001$). In addition, participants were 95.0% (± 1.4%) accurate on the distractor task. Behavioral results for Experiment 1 are depicted in Figure 2.

Lateralized ERP

In this experiment, we analyzed lateralized alpha power, the CDA, P_D, and N2pc. For all of the below analyses, we calculated the difference between contralateral and ipsilateral activity (see Fig. 3 for all neural results from Experiment 1, including average and standard error of the mean of the time course

Figure 2. Behavioral results from Experiment 1. Behavioral performance (K score) across the four conditions. Participants remembered fewer items when they were distracted than when they were not distracted. This impact of distraction was larger when participants had to discriminate the distractors than when they could ignore them. Average K score is represented by the horizontal black line and the black error bars reflect the standard error of the mean. The distribution of K scores in each condition for all participants is represented by the violin plots. Light gray lines connect data from one participant across conditions.

for each condition) and then took the difference (contralateral – ipsilateral). We then analyzed this difference value for the two conditions when distractors were presented (relevant vs. irrelevant) to determine whether there were any lateralized differences between conditions at each time window. To determine whether the signals were significantly lateralized, we additionally calculated one-way t-tests for each condition.

Predistractor (400–650 ms)
Before distractor onset (400–650 ms), there were no lateralized ERPs in either condition (Ignore: $t(19) = -1.78$, $P = 0.09$, $d = -0.40$; Discriminate: $t(19) = 0.51$, $P = -0.62$, $d = 0.11$), and there was no difference between trials with task-relevant versus task-irrelevant distractors ($t(19) = -1.53$, $P = 0.14$, $d = 0.24$). This is what we expected because the memory array was presented centrally and the lateralized distractors had not yet appeared.

Postdistractor I (850–950 ms)
During this time window (equivalent to 200–300 ms post-distraction, typically used for attention components), we observed a robust difference in lateralization between trials with task-relevant and task-irrelevant distractors ($t(19) = 3.19$, $P = 0.005$, $d = 0.71$). On trials with task-irrelevant distractors, the lateralized ERP was positive (M = 0.24 $\pm$ 0.45), and it was negative on trials with task-relevant distractors (M = -0.44 ± 0.98), suggesting the presence of an N2pc and P_D, respectively. To confirm this, we calculated the P_D and N2pc in specific a priori time windows (see Methods for detail). One-way, one-sample t-tests against zero confirmed that there was a reliable P_D when task-irrelevant distractors were presented ($t(19) = 3.19$, $P = 0.002$, $d = 0.71$) and no P_D when task-irrelevant distractors were presented ($t(19) = -1.65$, $P = 0.94$, $d = -0.37$). In addition, there
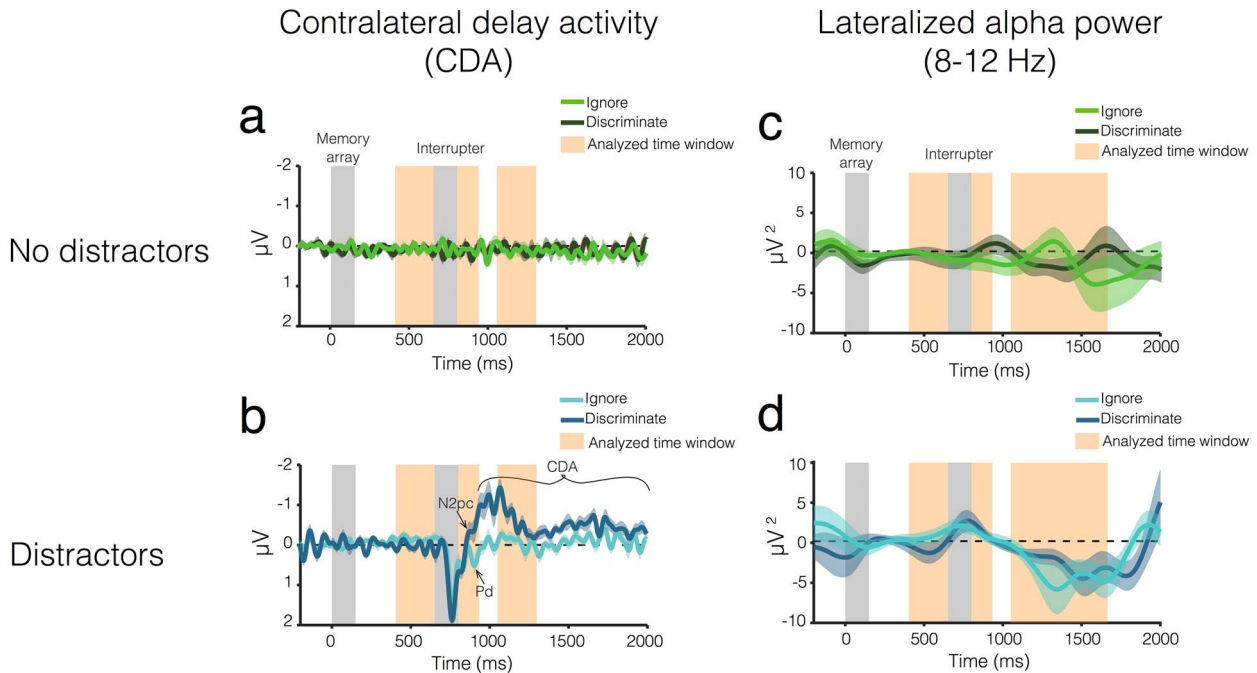
Figure 3. EEG results from Experiment 1. Average CDA amplitude over time for trials (*b*) with and (*a*) without distractors. The light color envelopes around each line represent standard error of the mean for each condition. The first vertical gray bar (0–150 ms) represents when the memory array was on the screen, and the second gray bar (650–800 ms) represents when the distractors were on the screen, if there were distractors on that trial. The orange vertical bars represent the analyzed time windows (400–650 ms; 850–950 ms; and 1050–1300 ms). Lateralized alpha power over time for trials (*d*) with and (*c*) without distractors.

was a reliable N2pc when task-relevant distractors were presented ($t(19) = -2.45$, $P = 0.01$, $d = -0.55$), but not when task-irrelevant distractors were presented ($t(19) = 2.37$, $P = 0.99$, $d = 0.53$).

Postdistractor II (1050–1300 ms)
In this time window, there was a significant difference in lateralization between trials with task-relevant and task-irrelevant distractors ($t(19) = 4.32$, $P < 0.001$, $d = 0.97$). The CDA was more lateralized when the task-relevant distractors were presented ($M = -0.59 \pm 0.55$) than when task-irrelevant distractors were presented ($M = -0.08 \pm 0.30$). In fact, the CDA was reliable on trials with task-relevant distractors (one-sample: $t(19) = -4.81$, $P < 0.001$, $d = -1.08$), but not on trials with task-irrelevant distractors (one-sample: $t(19) = -1.13$, $P = 0.27$, $d = -0.25$).

Lateralized Alpha Power

Predistractor (400–650 ms)
Before distractor onset (400–650 ms), alpha power was not significantly lateralized in either condition (Ignore: $t(19) = 1.13$, $P = 0.27$, $d = 0.25$; Discriminate: $t(19) = -0.57$, $P = 0.57$, $d = -0.13$), and there was no difference between trials with task-relevant versus task-irrelevant distractors ($t(19) = 0.991$, $P = 0.33$, $d = 0.22$). This was expected because the memory array was presented centrally, and the lateralized distractors had not yet appeared.

Postdistractor I (850–950 ms)
Immediately following distraction (850–950 ms), alpha power was not significantly lateralized in either condition (one-tailed *t*-tests against zero: Ignore: $t(19) = 1.20$, $P = 0.12$, $d = 0.27$;

Discriminate: $t(19) = 1.50$, $P = 0.08$, $d = 0.33$), and lateralization did not vary between conditions ($t(19) = -0.23$, $P = 0.82$, $d = -0.05$).

Postdistractor II (1050–1650 ms)
Towards the end of the trial (1050–1650 ms), alpha power was significantly lateralized in both conditions, consistent with a shift of spatial attention towards the distractors (one-tailed tests against zero: Ignore: $t(19) = -2.33$, $P = 0.015$, $d = -0.52$; Discriminate: $t(19) = -2.24$, $P = 0.019$, $d = -0.50$). Interestingly, there was no difference in lateralization between the two conditions (paired samples *t*-test: $t(19) = -0.89$, $P = 0.40$, $d = -0.20$).

Conclusions

In Experiment 1, participants performed a WM change detection task with distractors that appeared during the delay on a subset of trials. Behaviorally, participants remembered fewer items when they were distracted than when they were not distracted. This negative impact of distraction on behavior was larger when participants discriminated the task-relevant distractors than when they ignored task-irrelevant distractors.

The lateral position of the distractors allowed us to assess how they were processed using a suite of lateralized ERP signals. Task-relevant distractors elicited an N2pc followed by a sustained CDA when they had to be discriminated. This suggests that participants attend task-relevant distractors and then encode them into visual WM. Conversely, when task-irrelevant distractors were presented, there was a P_D instead of an N2pc and no CDA. Thus, task-irrelevant distractors were actively suppressed from being encoded into visual WM. In contrast, there is evidence that spatial attention may be captured regardless

of whether the distractors were task relevant, as shown by a decline in alpha power contralateral to the position of the distractors.

These findings suggest that observers could exert attentional control over whether the distractors entered into WM, and that this could be accomplished even when the distractors captured spatial attention. Thus, these findings converge with prior work that has pointed towards distinct computational roles for CDA and alpha activity, with the former associated with item-based storage, and the latter associated with covert spatial attention (Günseli et al. 2019; Hakim et al. 2019a).

Experiment 2

The relevance manipulation in Experiment 1 was a dual-task design, as it required participants to maintain information about two different tasks when task-relevant distractors were presented. With this kind of manipulation, participants could always try to optimize their performance by trading off between the two tasks on some portion of trials, especially during the relevant condition blocks. It is possible that participants simply chose to utilize an "offline" strategy on some trials in which they did not attempt to actively maintain the target items so that they could dedicate resources to the discrimination task. This strategy may be less likely in the irrelevant condition when subjects knew they do not need to do anything with the distractor items. Such a difference in strategy could plausibly explain why we observe a CDA to the task-relevant distractors and not for the task-irrelevant distractors. It is also generally consistent with our finding that the behavioral deficit was largest for the relevant condition. If this were the case, it would suggest that the results of Experiment 1 were the result of general strategic differences between the conditions that occur prior to distraction rather than the impact of the relevant distractors themselves. Therefore, in Experiment 2, we tested whether participants in the relevant condition actively encoded the target items into WM prior to the onset of the distractors, or whether they chose not to actively encode or maintain the target array in anticipation of making the discrimination. Therefore, the key question in Experiment 2 is whether there are differences in the CDA and alpha power lateralization between the ignore and discriminate conditions during the predistractor period. If participants did not actively store the memory array items in WM predistractor, the CDA should be reduced or eliminated in the relevant condition as compared with the irrelevant condition.

Materials and Methods

Participants

A total of 29 volunteers, naïve to the objective of the experiment, participated for payment ($15 USD per hour). The data from eight participants were excluded from the analysis because of too many artifacts (same criteria as Experiment 1). The remaining 21 participants (nine male) were between the ages of 30 and 18 (M = 22.1, SD = 3.9). We recruited participants from the same subject pool as Experiment 1, and we permitted participants from Experiment 1 to participate in Experiment 2. In total, four participants completed both experiments.

Apparatus and Stimuli

The apparatus was identical to Experiment 1. Stimuli (Fig. 4) were also identical to Experiment 1 with the following exceptions. We were interested in the neural representations of the memory array items. Therefore, we presented the memory array items laterally and the distractors centrally. Thus, CDA amplitude can be interpreted as encoding and maintenance of the memory array and lateralization of alpha power can be interpreted as a shift of attention towards laterally presented memory items. Therefore, at the beginning of the experiment, a horizontal diamond comprised of a green (RGB = 74, 183, 72; 52.8 cd/m^2) and a pink (RGB = 183, 73, 177; 31.7 cd/m^2) triangle appeared on the vertical midline 0.65° above the fixation dot. In 50% of the trials, the pink triangle pointed to the left side and the green triangle pointed to the right side, and in the remaining 50% of the trials, this was inverse. Half the participants were instructed to attend the hemifield that the pink triangle pointed to, and the other half was instructed to attend the hemifield to which the green triangle pointed. Memory displayed showed an array of three colored squares in each hemifield. Within each hemifield, there were one or two squares in the upper quadrant and two or one square in the lower quadrant. Squares could appear within an area of the display subtending 6° to the left or right of fixation and 3.1° above and below fixation. The distractor display showed four colored squares of the same size as the ones from the memory display along the midline of the screen, drawn from the remaining colors. These distracting items were shown on the vertical midline with a randomly jittered horizontal offset of maximally 0.55° (half of an object). Probe displays showed one colored square in each hemifield in the same location as one of the squares, randomly picked, in the original array. The color of the square in the unattended hemifield was the same as the original square on 50% of trials, and different on the other 50% of trials.
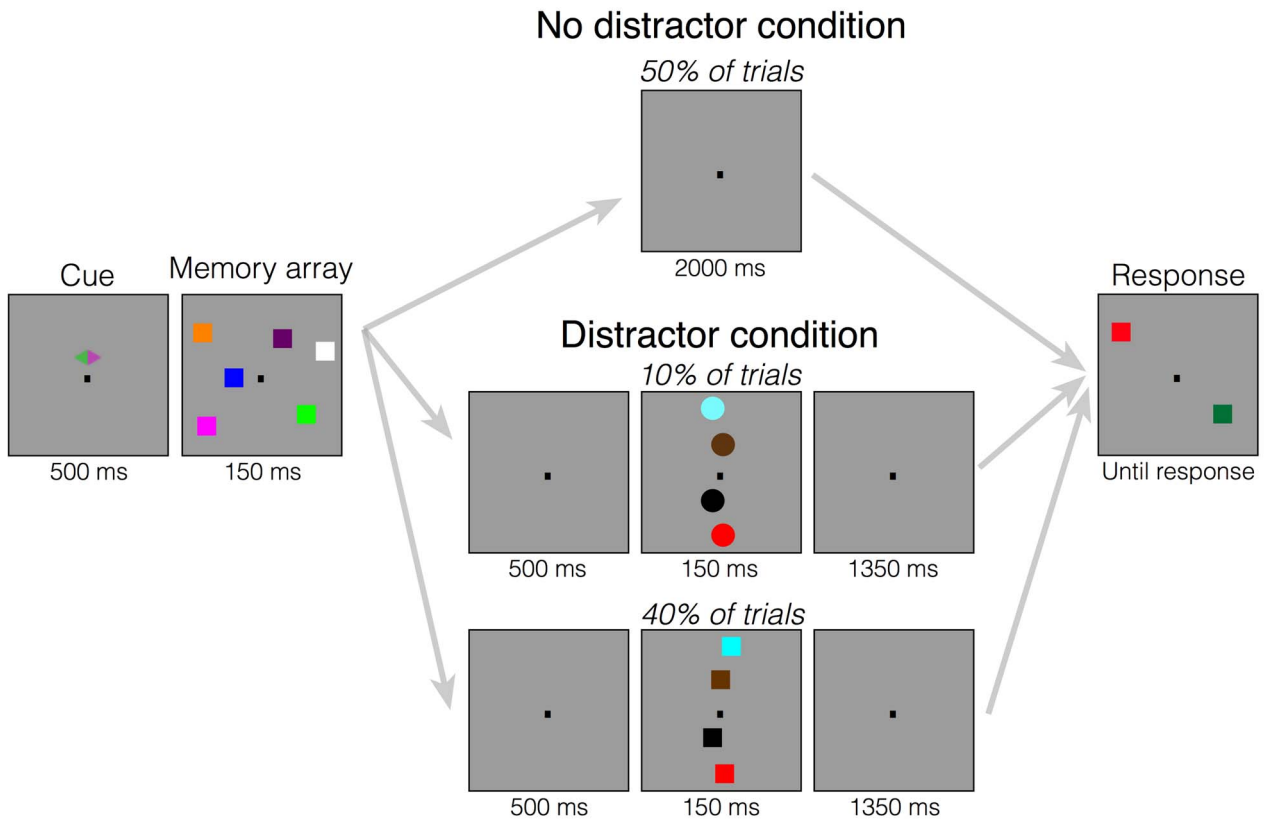
Procedure

The procedure was identical to Experiment 1 with the following exception. Each trial began with a cue display (500 ms) indicating the to-be-attended side of the screen (left or right).

Artifact Rejection and Analyses

Artifact rejection and analyses were identical to Experiment 1 with the following exceptions. Since circle distractors were only 10% of distractor trials and required an additional response, we only included square distractor trials in all analyses. In addition, we analyzed EEG data in all four conditions because stimuli were presented laterally in all cases. We compared conditions with a repeated measures ANOVA with the within-subjects factors Distraction (distractors vs. no distractors) and Relevance (ignore vs. discriminate). For alpha power, we analyzed all of the same time windows as the CDA because previous research has directly investigated the time course of alpha lateralization following centrally presented distractors (Hakim et al. 2020). Finally, we did not analyze the P$_D$ and N2pc because distractors were presented centrally in this experiment. We corrected all of the ANOVA results in the two postdistractor time windows for multiple comparisons using a Bonferroni correction for two comparisons (significance threshold: $P < 0.025$).

No distractor condition

50% of trials

2000 ms

Cue

500 ms

Memory array

150 ms

Distractor condition

10% of trials

500 ms   150 ms   1350 ms

40% of trials

500 ms   150 ms   1350 ms

Response

Until response

Figure 4. Task design for Experiment 2. At the start of each trial, a cue appeared above the fixation dot, which indicated to participants which side of the screen they should attend. Participants either attended to the green or purple side (counterbalanced across participants). Following the cue, the memory array appeared, which consisted of three colored squares on each side of the screen. Participants were told to remember the colors of the squares on the cued side. Following the memory array, the screen went blank. Then, either the screen remained blank the entire delay (no distractor conditions) or the screen went blank for 500 ms (distractor conditions) followed by a series of four objects (circles or squares) along the midline. When participants were in the "Ignore" condition, they were told to always ignore these distractor objects. When they were in the "Discriminate" condition, they were told to determine the shape of the stimuli in order to report whether the stimuli were circles. They were told to withhold their response until the response screen appeared. Following distractors, the screen then went blank for the rest of the delay. On the final screen, one square on either side of the screen reappeared and could either be the same color as the original square or it could be a different color that did not appear in the display. In both conditions, participants had to report whether the square on the cued side of the screen changed colors. In the "Discriminate" condition, participants additionally had to report whether the distractor objects were circles, if there were distractors on that trial.
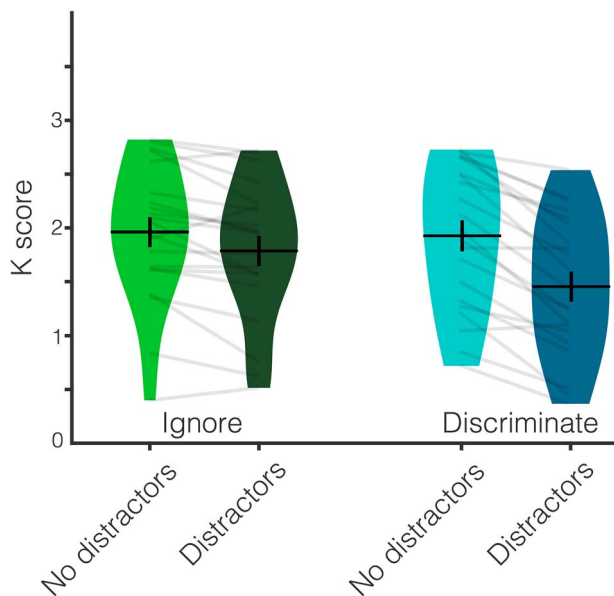
Results

Behavior

Behavioral performance was significantly above chance in all conditions (all one-sample t-tests, $P \leq 0.001$). Participants remembered fewer items when they were distracted (M $= 1.62 \pm 0.66$) than when they were not distracted (M $= 1.94 \pm 0.65$), significant main effect of Distraction ($F(1,20) = 100.21$, $P < 0.001$, $\eta_p^2 = 0.83$). Participants also remembered fewer items when they had to discriminate the distractors (M $= 1.69 \pm 0.64$) than when they could ignore the distractors (M $= 1.87 \pm 0.63$), significant main effect of Relevance ($F(1,20) = 10.20$, $P = 0.005$, $\eta_p^2 = 0.34$). The difference between trials with and without distractors was significantly larger when the task-relevant distractors had to be discriminated ($\Delta M = 0.479 \pm 0.221$) than when the task-irrelevant distractors could be ignored ($\Delta M = 0.177 \pm 0.258$), significant interaction of Distraction and Relevance ($F(1,20) = 13.67$, $P = 0.001$, $\eta_p^2 = 0.406$). The significant follow-up t-test showed that the difference between trials with and without

distractors was significantly larger in the Ignore than the Discriminate condition ($t(20) = -3.698$, $P = 0.001$). In addition, participants were 93.6% (± 3.5%) accurate on the distractor task. Behavioral results from Experiment 2 are depicted in Figure 5.

Lateralized ERP

Just as in Experiment 1, we ran a repeated measures ANOVA with the factors Distraction (no distraction, distraction) and Relevance (ignore, discriminate) to determine whether there were any differences between conditions at each time point. In this experiment, we analyzed lateralized alpha power and CDA. To determine whether the signals were significantly lateralized, we additionally calculated one-way t-tests for each condition. Results displayed in Figure 6 (figure includes the average and standard error of the mean of the time course for each condition).

Figure 5. Behavioral results from Experiment 2. Behavioral performance (K score) across the four conditions. Participants remembered fewer items when they were distracted than when they were not distracted. This impact of distraction was larger when participants had to discriminate the distractors than when they could ignore them. Average K score is represented by the horizontal black line and the black error bars reflect the standard error of the mean. The distribution of K scores in each condition for all participants is represented by the violin plots. Light gray lines connect data from one participant across conditions.

Predistractor (400–650 ms)
The CDA was reliable in all conditions (all one-sample t-tests, $P \leq 0.001$), and CDA amplitude did not vary between conditions (main effect of Distraction and Relevance and their interaction, all $P \geq 0.24$).

Postdistractor I (850–950 ms)
CDA amplitude was larger in the ignore condition (M = −0.90 ± 0.76) than in the discriminate (M = −0.59 ± 0.79), regardless of whether distractors appeared (significant main effect of Relevance: $F(1,20) = 2.08$, $P = 0.02$, $\eta_p^2 = 0.23$). The main effect of Distraction and the interaction of Distraction and Relevance were not significant (both $P \geq 0.12$).

Postdistractor II (1050–1300 ms)
CDA amplitude was larger on trials without distractors (M = −0.83 ± 0.51) than on trials with distractors (M = −0.22 ± 0.73), regardless of relevance (significant main effect of Distraction: $F(1,20) = 17.73$, $P < 0.001$, $\eta_p^2 = 0.47$). The main effect of Relevance was trending, but not significant ($F(1,20) = 3.65$, $P = 0.07$, $\eta_p^2 = 0.15$), and the interaction of Distraction and Relevance was not significant ($F(1,20) = 0.29$, $P = 0.69$, $\eta_p^2 = 0.01$). When distractors were not present, there was a reliable CDA regardless of condition (both $P \leq 0.001$).

Lateralized Alpha Power

Predistractor (400–650 ms)
Before the distractors appeared (400–650 ms), alpha power was significantly lateralized in all conditions (all $P \leq 0.004$), and lateralization did not vary between condition (main effect of Distraction and Relevance and their interaction were not significant, all $P \geq 0.25$).

Postdistractor I (850–950 ms)
Immediately following distractors (850–950 ms), alpha power was more lateralized on trials without distractors (M = −25.39 ± 33.33) than on trials with distractors (M = −3.71 ± 6.89), regardless of relevance (significant main effect of Distraction: $F(1,20) = 9.37$, $P = 0.006$, $\eta_p^2 = 0.32$). The main effect of Relevance and the interaction of Relevance and Distraction were not significant (both $P \geq 0.42$).

Postdistractor II (1050–1300 ms)
Toward the end of the delay (1050–1300 ms), alpha power was more lateralized on trials without distractors (M = −13.66 ± 24.62) than on trials with distractors (M = −3.77 ± 12.11; significant main effect of Distraction: $F(1,20) = 5.98$, $P = 0.02$, $\eta_p^2 = 0.23$), regardless of relevance. The main effect of Relevance and the interaction of Relevance and Distraction were not significant (both $P \geq 0.98$). Follow-up one-sample t-tests revealed that alpha power was significantly lateralized on trials without distractors (both $P \leq 0.03$), but not on trials with distractors (both $P \geq 0.09$), regardless of relevance. To summarize, the presentation of the central distractors disrupted alpha lateralization towards the target locations, and this effect did not depend on whether the distractors were task relevant.

Conclusions

In Experiment 2, we were interested in whether the relevance of the distractors affected the likelihood of encoding and maintaining target information in WM. The relevant distractor condition in our two experiments is similar to a dual-task design as participants have to maintain information about two separate tasks. Participants could decide to drop information about the initial target array in order to encode the new distracting information from the second task. Alternatively, they could attempt to sustain the initial WM representations at the expense of sufficiently attending the relevant distractors. How did participants address this trade-off between tasks in the present study?

Before the distractors appeared, there was no difference in CDA amplitude or alpha power lateralization between the ignore and discriminate conditions. There was a clear CDA and alpha power lateralization for both conditions, with no differences between them. This indicates that in both conditions, participants encoded and maintained lateralized WM representations and sustained their attention. Differences between conditions only emerged after the time when distractors were supposed to appear. Therefore, the results from Experiment 1 are not simply due to a dual-task trade-off between the encoding and maintenance of the memory array and the distractors.

The distractors always appeared at the same time during the delay. Therefore, participants could anticipate when they might be distracted. Around the time when distractors typically appeared, the CDA was smaller when participants anticipated a relevant distractor than when they anticipated a task-irrelevant distractor. During this same time window, alpha power lateralization depended on the presence of distractors. Towards the end of the trial, however, both CDA amplitude and alpha power lateralization were closer to baseline when distractors were present than when they were not, regardless of relevance. Overall, when participants anticipate that they may have to integrate new information into WM, they hold less information about the target array around the time that they think new information will appear, even if this new information does not

Contralateral delay activity (CDA)

Lateralized alpha power (8–12 Hz)
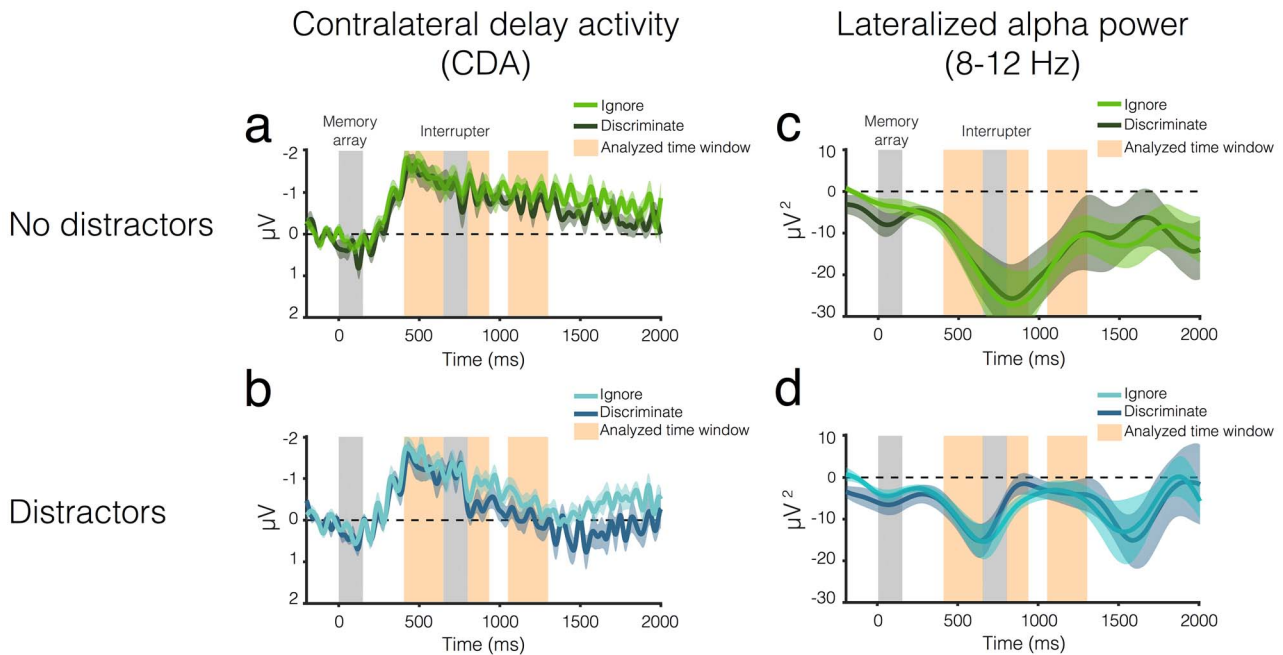
Figure 6. EEG results from Experiment 2. CDA amplitude over time for trials (*b*) with and (*a*) without distractors. The light color envelopes around each line represent standard error of the mean for each condition. The first vertical gray bar (0–150 ms) represents when the memory array was on the screen, and the second gray bar (650–800 ms) represents when the distractors were on the screen, if there were distractors on that trial. The orange vertical bars represent the analyzed time windows (400–650 ms; 850–950 ms; and 1050–1300 ms). Lateralized alpha power over time for trials (*d*) with and (*c*) without distractors.

appear. However, toward the end of the trial, relevance no longer impacts the likelihood of sustaining attention or maintaining information about the target array. Overall, we can rule out the alternate explanation of Experiment 1 that participants prematurely drop information about the target array when relevant distractors are presented. Participants were clearly attending and maintaining memoranda even when relevant distractors appeared.

Discussion

The key finding of the present study was that processing of stimuli that disrupt ongoing WM representations depended on their relevance. Task-relevant distractors were encoded and maintained in WM, whereas task-irrelevant distractors were suppressed and never entered WM. On the other hand, spatial attention was captured regardless of stimulus relevance. In Experiment 2, we investigated whether participants in the relevant condition of Experiment 1 actively encoded and maintained memory items prior to the onset of the distractors. We found that predistractor, there were no differences between conditions. Participants encoded and maintained WM representations and sustained attention to the memory items equally in both conditions. Differences between conditions only emerged after distractor onset, indicating that the results from Experiment 1 are not solely driven by a dual-task trade-off between the maintenance of the memory array and the distractors.

We observed distinct effects of distraction on lateralized ERP signals and lateralized alpha power in these experiments. The results converge with past proposals of a distinction between item-based and spatial capture of attention (Hakim et al. 2019a,

2020). In our procedure, alpha oscillations showed that distractors captured spatial attention, regardless of task relevance. By contrast, the formation of item-based representations in WM was completely determined by task relevance, such that encoding into WM was suppressed when observers could ignore the distractors. Interestingly, both item-based storage and spatial attention towards the memoranda were eventually disrupted by the presentation of the distractors, regardless of whether a dual task was imposed.

Spatial Capture by Task-Relevant and Task-Irrelevant Distractors

In Experiment 1, we directly investigated how distractors are processed. In this experiment, we found that alpha power was significantly lateralized following the onset of both task-relevant and task-irrelevant distractors. This result provides evidence that spatial attention shifted to the location of both the relevant and irrelevant distractors. We designed our distractors to be very similar to the original memoranda in order to induce a large behavioral deficit on distractor trials compared with nondistractor trials. Previous research has shown that when distractors share features with memoranda (Soto et al. 2008; Olivers and Eimer 2011; van Moorselaar et al. 2015; Hollingworth and Beck 2016) or are part of the attentional set (Folk and Remington 1998; Folk et al. 2008), spatial attention can be captured at the location of the distractors. However, the similarity between targets and distractors is a continuum (Duncan and Humphreys 1989) and attentional capture increases with target–distractor similarity (Ludwig and Gilchrist 2002; Ansorge and Heumann 2003). In the present study, relevant distractors were the same shape as the targets and drew

their color from the same pool of colors as the targets, that is, they had potential target colors. However, no distractors were ever the same color as a target within a trial. Nevertheless, the finding that spatial attention was captured by both types of distractors may be partially due to the fact that even irrelevant distractors were perceptually similar enough to the targets (i.e., contingent capture). Future research should determine whether spatial attention is necessarily captured when distractors do not share any similarities with currently maintained WM representations. Understanding whether spatial and item-based attention are captured along the entire continuum is an important question as it will give insight into the question of how features are weighted in WM and how this affects attention deployment.

Voluntary Control of Item-Based Capture

In Experiment 1, irrelevant distractors elicited a P_D, suggesting that they were actively suppressed. Relevant distractors, however, first elicited an N2pc, which indicates that these items were individuated. The subsequent lateralized negativity then transitioned into a CDA, suggesting that distractors were not only individuated but also encoded into WM. These results illustrate that participants dynamically respond to task demands by suppressing irrelevant distractors from WM and only encoding relevant distractors into WM. Even when salient stimuli captured spatial attention, participants still had voluntary control over whether to store that information in WM. These results are in line with previous findings showing that successful suppression of irrelevant information can contribute to better performance (Sawaki and Luck 2012; Gaspar and McDonald 2014; Feldmann-Wüstefeld et al. 2016; Weaver et al. 2017). For example, when target identity is correctly reported in a visual search task, a concurrently presented salient distractor elicits a pronounced P_D component, indicative of active suppression (Feldmann-Wüstefeld et al. 2020). Conversely, when the distractor identity is erroneously reported, distractors elicit a CDA and a less pronounced P_D, suggesting that the distractor was encoded into WM.

Our results also nicely align with the WM gating literature. This literature provides a framework to explain which information is allowed to enter WM and which information is blocked (O'Reilly and Frank 2006; Badre 2012; Chatham et al. 2014; Chatham and Badre 2015). According to this account, the WM gate is the mechanism by which irrelevant information is blocked from entering. When the WM gate is open, it allows information to enter WM. When it is closed, ongoing WM representations are sustained, whereas irrelevant information is blocked (Badre 2012). The WM gating literature has mostly used functional magnetic resonance imaging to demonstrate which parts of the brain are involved in WM gating and maintenance and has not distinguished between WM gating and capture of spatial attention. Our results suggest that we may be able use EEG activity to track WM gating as well. We propose that the CDA tracks how much information passes through the gate, whereas the P_D reflects the gate itself. Previous research has shown that P_D amplitude scales with the number of items that were blocked from entering WM (Feldmann-Wüstefeld et al. 2019). Therefore, P_D amplitude may reflect how firmly the WM gate was closed. Conversely, the CDA could reflect how much information is encoded into WM, with relevant information more likely to pass through the gate. Our results suggest that how firmly the gate is closed can be controlled in a top-down

manner, whereas other factors, such as physical salience, determine how much information enter WM. Future research could investigate the precise temporal dynamics of WM input and output gating using these proposed EEG signals. For example, if the WM gate accidentally allows irrelevant information into WM, how long is that information maintained in WM before it is dropped?

Is Attentional Capture Obligatory?

Our findings provide a new framework in which we can investigate attentional capture. We propose that attentional capture is comprised of item-based and spatial capture. Item-based capture involves forming WM representations of the new stimuli. Based on our findings, item-based capture appears to be subject to voluntary attentional control. It allows relevant stimuli to enter WM, whereas irrelevant stimuli are suppressed. We also have clear evidence that, in our specific task context, spatial capture occurred when distractors were present, regardless of top-down goals. However, whether spatial capture merely happened because of perceptual salience (i.e., bottom-up capture) or whether similarity with task-relevant items (continent capture) contributed to attention deployment is unclear and should be the focus of future research. Thus, WM gating can successfully block irrelevant information, even when spatial attention is captured. From our perspective, information needs to be held in WM in order for it to be processed. Therefore, if distracting stimuli are not encoded into WM via item-based capture, they are not fully processed, even if they captured spatial attention. Selecting an item using spatial attention is necessary, but not sufficient, to encode it into WM.

We investigated item-based and spatial capture during maintenance of ongoing WM representations. However, the majority of the attentional capture literature has investigated this process during encoding. Therefore, future research could apply this new framework of attentional capture to investigate item-based and spatial capture during encoding. This could potentially provide insight into the ongoing debate about whether attentional capture during encoding is obligatory (Hickey et al. 2009; Sawaki et al. 2012; Feldmann-Wüstefeld and Schubö 2013; Gaspar and McDonald 2014; Gaspelin et al. 2015, 2017; Liesefeld et al. 2017; Gaspelin and Luck 2018; Feldmann-Wüstefeld et al. 2020). We hypothesize that the ways in which the two subcomponents of attentional capture (item-based and spatial capture) respond during encoding should be similar to how they respond during maintenance. That is, we hypothesize that during encoding, spatial capture may happen regardless their relevance, whereas item-based capture may be subject to voluntary attentional control.

To further elucidate the neural mechanism underlying WM performance, future research could also compare the EEG signal for accurate versus inaccurate responses. In the current study, behavioral performance was very high (mean accuracy of 79.9% and 78.7% in Experiments 1 and 2), which means that an insufficient number of trials was available for a reliable comparison. A direct comparison of correct and incorrect trials could reveal which cognitive mechanism is involved in WM failures. For example, previous research found that the P_D component is smaller on incorrect than on correct trials in a visual search task (Feldmann-Wüstefeld et al. 2019). For the present research question, it could be particularly insightful to compare CDA and alpha-band activity between correct and incorrect trials to reveal whether item-based or spatial attention contribute more to behavioral errors.

Impact of Distractors on Ongoing WM Representations

Previous research has demonstrated that salient distractor stimuli interfere with object representations in WM and cause attention to shift away from maintained representations (Hakim et al. 2019a). Our work replicates and extends these findings by adding a top-down perspective. Participants had to either attend (relevant) or ignore (irrelevant) distractors. Here, we show that the CDA, a neural measure of WM load, was initially influenced by top-down goals. When participants anticipated that they may have to encode additional information into WM, the CDA was smaller than when participants knew that distractor stimuli could be ignored. However, the CDA was at baseline toward the end of trials that contained both relevant and irrelevant distractors. This suggests that both types of distractors harmed lateralized object representations of the memoranda. However, towards the end of the trial, alpha power, a neural measure of spatial attention, shifted to baseline following both relevant and irrelevant distractors, suggesting that participants shifted their attention away from the locations of the original memoranda, regardless of whether the distractor information was relevant. These results illustrate that the observers' goals determine the encoding of item-based representations, even when spatial attention is captured. Toward the end of the trial, however, distractors harm both spatial attention and the object representations of the memoranda, regardless of top-down goals.

Subcomponent Processes of Attentional Capture

Our two experiments provide evidence that attentional capture may be comprised of at least two distinct subcomponent processes: item-based capture and spatial capture. Future research should investigate how these subprocesses interact with other forms of attention, such as sustained attention and vigilance. For example, previous research has found that participants with lower executive WM capacity tend to mind-wander more than those with higher executive WM capacity (Kane et al. 2007; McVay and Kane 2009, 2010). Does an analogous relationship exist between spatial attention and mind-wandering? In addition, if a participant is in a high attentional state, is their spatial attention less likely to be captured? Thus far, research on sustained attention, spatial attention, and mind-wandering have largely developed in parallel. The type of experimental design that we used in the current study may provide a fruitful avenue to integrate these currently disparate views of attention.

Conclusions

Previous research has treated attentional capture as a monolithic process. Here, we present new evidence that there are at least two subcomponent processes of attentional capture that are neurally dissociable: spatial capture and item-based capture. Lateralized alpha power indexes spatial capture, a process that involves a shift of spatial attention. By contrast, item-based capture is tracked by the N2pc and CDA when item-based representations are deemed relevant and allowed to enter WM, whereas the P_D tracks the active suppression of items from WM. This fractionation of attentional capture into distinct subcomponent processes provides a framework by which the fate of ongoing WM processes after distractors can be explained. We show that relevant distractors trigger both of these dissociable processes. Irrelevant distractors, however, only trigger spatial capture.

Supplementary Material

Supplementary material can be found at *Cerebral* Cortex online.

Authors' Contributions

N.H., T.F.W., and E.K.V. conceived the study. N.H. and T.F.W. performed analyses. N.H. collected the data and wrote the initial draft of the manuscript, which all authors read and edited.

Data Availability

Datasets for all experiments will become available online on Open Science Framework upon acceptance of the manuscript or reviewer request.

Notes

Conflict of Interest: None declared.

References

Ansorge U, Heumann M. 2003. Top-down contingencies in peripheral cuing: the roles of color and location. *J Exp Psychol Hum Percept Perform*. 29(5):937–948. doi: 10.1037/0096-1523.29.5.937.

Awh E, Belopolsky AV, Theeuwes J. 2012. Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in cognitive sciences*. 16(8):437–443.

Badre D. 2012. Opening the gate to working memory. *Proc Natl Acad Sci*. 109(49):19878–19879. doi: 10.1073/pnas.1216902109.

Balaban H, Luria R. 2017. Neural and behavioral evidence for an online resetting process in visual working memory. *J Neurosci*. 37(5):1225–1239. doi: 10.1523/JNEUROSCI.2789-16.2016.

Bisley JW, Goldberg ME. 2010. Attention, intention, and priority in the parietal lobe. *Annu Rev Neurosci*. 33(1):1–21. doi: 10.1146/annurev-neuro-060909-152823.

Bundesen C, Habekost T, Kyllingsbæk S. 2005. A neural theory of visual attention: bridging cognition and neurophysiology. *Psychological review*. 112(2):291.

Bundesen C. 1990. A theory of visual attention. *Psychological review*. 97(4):523.

Burra N, Kerzel D. 2014. The distractor positivity (Pd) signals lowering of attentional priority: evidence from event-related potentials and individual differences. *Psychophysiology*. 51(7):685–696. doi: 10.1111/psyp.12215.

Chatham CH, Badre D. 2015. Multiple gates on working memory. *Curr Opin Behav Sci*. 1:23–31. doi: 10.1016/j.cobeha.2014.08.001.

Chatham CH, Frank MJ, Badre D. 2014. Corticostriatal output gating during selection from working memory. *Neuron*. 81(4):930–942.

Clapp WC, Gazzaley A. 2012. Distinct mechanisms for the impact of distraction and interruption on working memory in aging. *Neurobiol Aging*. 33(1):134–148. doi: 10.1016/j.neurobiolaging.2010.01.012.

Corbetta M, Shulman GL. 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*. 3(3):201–215.

Cowan N. 2001. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and brain sciences*. 24(1):87–114.

Desimone, R, Duncan, J. (1995). Neural mechanisms of selective visual attention. Annual review of neuroscience, 18(1), 193–222.

Duncan J, Humphreys GW. 1989. Visual search and stimulus similarity. *Psychological review*. 96(3):433.

Egeth HE, Yantis S. 1997. Visual attention: control, representation, and time course. *Annual review of psychology*. 48(1): 269–297.

Feldmann-Wüstefeld T, Brandhofer R, Schubö A. 2016. Rewarded visual items capture attention only in heterogeneous contexts. *Psychophysiology*. 53(7):1063–1073. doi: 10.1111/psyp.12641.

Feldmann-Wüstefeld T, Vogel EK. 2019. Neural evidence for the contribution of active suppression during working memory filtering. *Cerebral Cortex*. 29(2):529–543.

Feldmann-Wüstefeld T, Busch NA, Schubö A. 2019. Failed suppression of salient stimuli precedes behavioral errors. *J Cogn Neurosci*. 32(2):367–377. doi: 10.1162/jocn_a_01502.

Feldmann-Wüstefeld T, Schubö A. 2013. Context homogeneity facilitates both distractor inhibition and target enhancement. *J Vis*. 13(2013):1–12. doi: 10.1167/13.3.11.doi.

Feldmann-Wüstefeld T, Uengoer M, Schubö A. 2015. You see what you have learned. Evidence for an interrelation of associative learning and visual selective attention. *Psychophysiology*. 52(11):1483–1497. doi: 10.1111/psyp.12514.

Feldmann-Wüstefeld T, Vogel EK, Awh E. 2018. Contralateral delay activity indexes working memory storage, not the current focus of spatial attention. *Journal of cognitive neuroscience*. 30(8):1185–1196.

Folk CL, Leber AB, Egeth HE. 2008. Top-down control settings and the attentional blink: evidence for nonspatial contingent capture. *Vis Cogn*. 16(5):616–642. doi: 10.1080/13506280601134018.

Folk CL, Remington RW, Johnston JC. 1992. Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human perception and performance*. 18(4):1030.

Folk CL, Remington R. 1998. Selectivity in distraction by irrelevant featural singletons: evidence for two forms of attentional capture. *J Exp Psychol Hum Percept Perform*. 24(3):847–858. doi: 10.1037/0096-1523.24.3.847.

Foster JJ, Bsales EM, Jaffe RJ, Awh E. 2017. Alpha-band activity reveals spontaneous representations of spatial position in visual working memory. *Curr Biol*. 27(20):3216–3223.e6. doi: 10.1016/j.cub.2017.09.031.

Foster JJ, Sutterer DW, Serences JT, Vogel EK, Awh E. 2016. The topography of alpha-band activity tracks the content of spatial working memory. *Journal of neurophysiology*. 115(1):168–177.

Franconeri SL, Simons DJ. 2003. Moving and looming stimuli capture attention. *Perception & psychophysics*. 65(7):999–1010.

Gaspar JM, McDonald JJ. 2014. Suppression of salient objects prevents distraction in visual search. *J Neurosci*. 34(16):5658–5666. doi: 10.1523/JNEUROSCI.4161-13.2014.

Gaspelin N, Leonard CJ, Luck SJ. 2015. Direct evidence for active suppression of salient-but-irrelevant sensory inputs. *Psychol Sci*. 26(11):1740–1750. doi: 10.1177/0956797615597913.

Gaspelin N, Leonard CJ, Luck SJ. 2017. Suppression of overt attentional capture by salient-but-irrelevant color singletons. *Attention, Perception, & Psychophysics*. 79(1):45–62.

Gaspelin N, Luck SJ. 2018. Combined electrophysiological and behavioral evidence for the suppression of salient distractors. *J Cogn Neurosci*. 30(9):1265–1280. doi: 10.1162/jocn_a_01279.

Günseli E, Fahrenfort JJ, van Moorselaar D, Daoultzis KC, Meeter M, Olivers CN. 2019. EEG dynamics reveal a dissociation between storage and selective attention within working memory. *Scientific reports*. 9(1):1–13.

Hakim N, Adam KCS, Gunseli E, Awh E, Vogel EK. 2019a. Dissecting the neural focus of attention reveals distinct processes for spatial attention and object-based storage in visual working memory. *Psychol Sci*. 30(4):526–540. doi: 10.1177/0956797619830384.

Hakim N, Awh E, Vogel EK. 2020. Manifold visual working memory. Working memory: the state of the. *Science*. 311.

Hakim N, Feldmann-Wüstefeld T, Awh E, Vogel EK. 2019b. Perturbing neural representations of working memory with task-irrelevant interruption. *J Cogn Neurosci*. 32(3):558–569. doi: 10.1162/jocn_a_01481.

Hickey C, Di Lollo V, McDonald JJ. 2009. Electrophysiological indices of target and distractor processing in visual search. *J Cogn Neurosci*. 21(4):760–775. doi: 10.1162/jocn.2009.21039.

Hollingworth A, Beck VM. 2016. Memory-based attention capture when multiple items are maintained in visual working memory. *J Exp Psychol Hum Percept Perform*. 42(7):911–917. doi: 10.1037/xhp0000230.

Itti L, Koch C. 2000. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision research*. 40(10–12):1489–1506.

Jonides J, Irwin DE. 1981. *Capturing attention*.

Kane MJ, Brown LH, McVay JC, Silvia PJ, Myin-Germeys I, Kwapil TR. 2007. For whom the mind wanders, and when: an experience-sampling study of working memory and executive control in daily life. *Psychol Sci*. 18(7):614–621. doi: 10.1111/j.1467-9280.2007.01948.x.

Kastner S, Ungerleider LG. 2001. The neural basis of biased competition in human visual cortex. *Neuropsychologia*. 39(12):1263–1276.

Liesefeld HR, Liesefeld AM, Töllner T, Müller HJ. 2017. Attentional capture in visual search: capture and post-capture dynamics revealed by EEG. *Neuroimage*. 156:166–173. doi: 10.1016/j.neuroimage.2017.05.016.

Ludwig CJH, Gilchrist ID. 2002. Stimulus-driven and goal-driven control over visual selection. *J Exp Psychol Hum Percept Perform*. 28(4):902–912. doi: 10.1037/0096-1523.28.4.902.

Luria R, Balaban H, Awh E, Vogel EK. 2016. The contralateral delay activity as a neural measure of visual working memory. *Neurosci Biobehav Rev*. 62:100–108. doi: 10.1016/j.neubiorev.2016.01.003.

Maxwell JW, Gaspelin N, Ruthruff E. 2020. No identification of abrupt onsets that capture attention: evidence against a unified model of spatial attention. *Psychological Research*. 1–17.

Mazza V, Turatto M, Umiltà C, Eimer M. 2007. Attentional selection and identification of visual objects are reflected by distinct electrophysiological responses. *Exp Brain Res*. 181(3):531–536. doi: 10.1007/s00221-007-1002-4.

Mcdonald JJ, Green JJ, Jannati A, Di Lollo V. 2013. On the electrophysiological evidence for the capture of visual attention. *J Exp Psychol Hum Percept Perform*. 39(3):849–860. doi: 10.1037/a0030510.

McVay JC, Kane MJ. 2009. Conducting the train of thought: working memory capacity, goal neglect, and mind wandering

in an executive-control task. *J Exp Psychol Learn Mem Cogn*. 35(1):196–204. doi: 10.1037/a0014104.

McVay JC, Kane MJ. 2010. Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychol Bull*. 136(2):188–207. doi: 10.1037/a0018298.

Olivers CNL, Eimer M. 2011. On the difference between working memory and attentional set. *Neuropsychologia*. 49(6):1553–1558. doi: 10.1016/j.neuropsychologia.2010.11.033.

Olivers CNL, Meijer F, Theeuwes J. 2006. Feature-based memory-driven attentional capture: visual working memory content affects visual attention. *J Exp Psychol Hum Percept Perform*. 32(5):1243–1265. doi: 10.1037/0096-1523.32.5.1243.

O'Reilly RC, Frank MJ. 2006. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput*. 18(2):283–328. doi: 10.1162/089976606775093909.

Posner MI, Snyder CR, Solso R. 2004. Attention and cognitive control. *Cognitive psychology: Key readings*. 205.

Posner MI, Petersen SE. 1990. The attention system of the human brain. *Annual review of neuroscience*. 13(1):25–42.

Prinzmetal W, Zvinyatskovskiy A, Gutierrez P, Dilem L. 2009. Voluntary and involuntary attention have different consequences: the effect of perceptual difficulty. *Q J Exp Psychol*. 62(2):352–369. doi: 10.1080/17470210801954892.

Sawaki R, Geng JJ, Luck SJ. 2012. A common neural mechanism for preventing and terminating the allocation of attention. *J Neurosci*. 32(31):10725–10736. doi: 10.1523/JNEUROSCI.1864-12.2012.

Sawaki R, Luck SJ. 2012. Active suppression of distractors that match the contents of visual working memory. *Vis Cogn*. 19(7):1–14. doi: 10.1080/13506285.2011.603709.Active.

Soto D, Hodsoll J, Rotshtein P, Humphreys GW. 2008. Automatic guidance of attention from working memory. *Trends Cogn Sci*. 12(9):342–348. doi: 10.1016/j.tics.2008.05.007.

Spinks JA, Zhang JX, Fox PT, Gao J-H, Hai Tan L. 2004. More workload on the central executive of working memory, less attention capture by novel visual distractors: evidence from an fMRI study. *Neuroimage*. 23(2):517–524. doi: 10.1016/j.neuroimage.2004.06.025.

Theeuwes J. 2010. Top–down and bottom–up control of visual selection. *Acta Psychol (Amst)*. 135(2):77–99. doi: 10.1016/j.actpsy.2010.02.006.

Trafton JG, Altmann EM, Brock DP, Mintz FE. 2003. Preparing to resume an interrupted task: effects of prospective goal encoding and retrospective rehearsal. *Int J Hum Comput Stud*. 58(5):583–603. doi: 10.1016/S1071-5819(03)00023-5.

van Moorselaar D, Gunseli E, Theeuwes J, NL Olivers C. 2015. The time course of protecting a visual memory representation from perceptual interference. *Frontiers in human neuroscience*. 8:1053.

van Moorselaar D, Gunseli E, Theeuwes J, Olivers CNL. 2015. The time course of protecting a visual memory representation from perceptual interference. *Front Hum Neurosci*. 8. doi: 10.3389/fnhum.2014.01053.

Weaver MD, Hickey C, van Zoest W. 2017. The impact of salience and visual working memory on the monitoring and control of saccadic behavior: an eye-tracking and EEG study. *Psychophysiology*. 54(4):544–554. doi: 10.1111/psyp.12817.

Wolfe JM, Horowitz TS. 2004. What attributes guide the deployment of visual attention and how do they do it? *Nature reviews neuroscience*. 5(6):495–501.

Yantis S, Jonides J. 1984. Abrupt visual onsets and selective attention: evidence from visual search. *Journal of Experimental Psychology: Human perception and performance*. 10(5):601.

Zivony A, Allon AS, Luria R, Lamy D. 2018. Dissociating between the N2pc and attentional shifting: an attentional blink study. *Neuropsychologia*. 121:153–163.